Strength in Numbers

Remember, no matter how well-trained an individual model is, there is still a significant chance that it could perform poorly or produce incorrect results. Rather than relying on a single model, you can often get better results by training multiple models or using multiple algorithms and in some way capturing the collective results. As we mentioned, there are two main approaches to this: **Ensemble learning** and **automated machine learning**. Let's have a closer look at each of them.

Ensemble Learning

Remember, **ensemble learning** combines multiple machine learning models to produce one predictive model. There are three main types of ensemble algorithms:

Bagging or **bootstrap aggregation**

- Helps reduce overfitting for models that tend to have high variance (such as *decision trees*)
- Uses random subsampling of the training data to produce a *bag* of trained models.
- The resulting trained models are homogeneous
- The final prediction is an average prediction from individual models

Boosting

- Helps reduce bias for models.
- In contrast to bagging, boosting uses the same input data to train multiple models using different hyperparameters.
- Boosting trains model in *sequence* by training weak learners one by one, with each new learner correcting errors from previous learners
- The final predictions are a *weighted* average from the individual models

Stacking

- Trains a large number of completely different (heterogeneous) models
- Combines the outputs of the individual models into a meta-model that yields more accurate predictions

Strength in Variety: Automated ML

Automated machine learning, like the name suggests, automates many of the iterative, time-consuming, tasks involved in model development (such as selecting the best features, scaling features optimally, choosing the best algorithms, and tuning hyperparameters). Automated ML allows data scientists, analysts, and developers to build models with greater scale, efficiency, and productivity—all while sustaining model quality.

QUIZ QUESTION

Luis has been experimenting with a machine learning algorithm that make predictions by calculating the weighted averages of weak classifiers. What is the type of machine learning algorithm Luis is working with?

- ○ Bagging
- ✓ Boosting
- ○ Stacking

SUBMIT

Now that we've talked about the concepts underlying ensemble learning and automated machine learning, let's get some hands-on practice with both of these approaches in Azure Machine Leaning Studio.

NEXT